ANFIS Based Intelligent Solar Flare Prediction System

Hussain Ageel Naseer[1], Ajmal Hussain Shah[2*], Raed Abdulla[3]
[1,2,3]School of Engineering
Asia Pacific University of Technology and Innovation Technology Park Malaysia, Bukit Jalil, Malaysia
* KICT, International Islamic University Malaysia (IIUM), Selangor, 53100 Malaysia
agyl@hotmail.com, ajmal@apu.edu.my, raed@apu.edu.my

ABSTRACT

Satellite plays vital role in conveying information around the world. Satellites communication systems provide various services such as TV broadcasting, internet, telephone, telemedicine and GPS. However, these systems are threatened by solar radiation so called solar flare erupted from the Sun which is one of the major threats faced by satellites orbiting around the Earth. Thus, to keep satellites from being damaged an intelligent ANFIS based solar flare prediction system is proposed in this paper. The system is trained with publicly available time series data of flares occurred in the past from 2001-2011. The results show that the system has the accuracy of 85.47 % and it is capable of detecting the flare 24 hour before its occurrence.

KEYWORDS

Solar Flare, flare prediction, Artificial Neuro-Fuzzy Inference System (ANFIS), Membership Functions, Neural Network.

1 INTRODUCTION

Telecommunication industry is one of the fastest growing sectors in engineering, which conveys information over a distance by means of electrical or electromagnetic signals. Beacons signal flags and telegraphs are some of the applications of telecommunication before it has adopted wireless technologies.

Radio (or wireless) communication shares information between two points via free-space by means of electromagnetic waves. A transmitter, medium (free space) and a receiver are the three main components of radio communication system. One of the distinct features of a radio communication system is that it does not need solid medium like copper cable which significantly minimizes the cost. In addition, it requires a good line of sight to transmit data up to 70 km. On the other hand, satellite communication covers large area and only 3 GEO satellites are required for global coverage.

A satellite is an object orbiting around a planet or star. A communication satellite is an artificial object orbiting around the Earth which consists of a transmitter and a receiver in order to repeat and/or modulate a received signal and transmit back to desired destination [1]. Mobile communication system, television broadcasting, and Global Positioning System (GPS) services are some of the commonly used applications of satellite today. Based on the applications, the satellites are orbiting at an altitude of 2000 km to 35800 km.

The satellites, however, are threatened by various factors. Temperature, debris, and solar flares are some of the common threats to satellites. Satellites are located above the Earth surface at altitudes outside the Ozone. Satellite on-board systems operate on electrical energy which is gained by solar cells. Direct sunlight with high temperatures fallen on satellite causes malfunctioning of the on-board systems. This results in weakening the cement safeguarding the solar cells, decreasing the power output from them.

Similarly, extreme cold in surface due to darkness may also cause damage to satellites [2].
A solar flare is high energy burst erupted from the Sun. solar flares are caused due to the sudden change in the magnetic field (due to x-ray present) around the surface of the sun. The solar flares are classified based on the concentration of X-ray within them: C, M, and X class flares are considered harmful, in which X class flare consists of the highest concentration of X-ray. Due to the high concentration of X-ray in X class flare, it results in solar storms for longer period of time than the others. [3-4].

Likewise a solar flare is the most powerful solar activity which determines the space weather. Unlike the previous classification, a flare cannot be at predicted at all times based on the electromagnetic status of the flares. It presents a technique whereby the flares are classified based on solar image observations [5].

Solar radiation is one of the major threats faced by the satellites orbiting around the earth. Solar flares are the intense blooms of radiation which are released from the sun due to magnetic storms on the surface of the sun (sunspots). These flares are considered as a threat due to its magnetic characteristics which could result in communication outage, or worst case scenario: satellite breakdown.

2 DATA

Solar flares are grouped either in McIntosh or Mt. Wilson classification. The different prediction systems select either McIntosh or Mt. Wilson classification. Also large-scale images were used by some researchers claiming that it can provide more detailed information about the flares.

The McIntosh classification is based on the intensity of X-ray produced by the solar flare, while Mt. Wilson is based on the magnetic characteristics of the solar flare. It is noticed that majority of the researchers use flared data on McIntosh classification. In addition, McIntosh classification is a standard for international exchange of solar geophysics, which is also publicly available. More importantly, the many research papers have adopted McIntosh classification in their experiments, thus use of another classification would not provide a fair comparison. In this paper, McIntosh classification based solar flare time series data has been used. Two thousand data-pairs were collected for past 10 years from 2001 to 2011. Out of which, half of the pairs were used for training the system, while the other half were used for testing/validating the performance of the proposed system. The data is publicly available on NGDC website. A sample of the solar flare data is shown in Figure 1. The figure shows Event ID, Date, Start and End time and GOES class of flares.

EventID	Date	Start (UT)	Max (UT)	End (UT)	Pos.	GOES Class	Keyword
20061214_2215	14-Dec-06	21:07:00	22:15:00	22:26:00	S05W31	X1.5	Flare
20061213_0240	13-Dec-06	02:14:00	02:40:00	02:57:00	S06W22	X3.4	Flare
20061207_1913	07-Dec-06	18:20:00	19:13:00	19:33:00		M2.0	Flare
20061206_2019	06-Dec-06	20:14:00	20:19:00	20:22:00		M3.5	Flare
20061206_1847	06-Dec-06	18:29:00	18:47:00	19:00:00		X6.5	Flare
20061206_0823	06-Dec-06	08:02:00	08:23:00	09:03:00		M6.0	Flare
20061206_0220	06-Dec-06	01:30:00	02:20:00	02:54:00	S07E69	M1.1	Flare
20061205_1035	05-Dec-06	10:18:00	10:35:00	10:45:00		X9.0	Flare
20061205_0803	05-Dec-06	07:45:00	08:03:00	08:06:00		M1.8	Flare
20060706_0836	06-Jul-06	08:13:00	08:36:00	08:51:00		M2.5	Flare
20060427_1552	27-Apr-06	15:22:00	15:52:00	15:58:00		M7.9	Flare
20060426_1702	26-Apr-06	16:51:00	17:02:00	17:10:00		M1.3	Flare

Figure 1. Sample of raw data collected from NGDC.

3 DESIGN, IMPLEMENTATION AND TESTING

The solar flare prediction is devised using the concept of time series prediction and its performance has been analyzed. The construction details and the design process of the flare prediction system are discussed in the following sections.

3.1 Artificial Neuro-Fuzzy Inference System (ANFIS)

ANFIS is a hybrid system which is a combination of neural network and fuzzy logic. It is a hybrid leaning algorithm, combination of back-propagation and least square estimation is adopted in system modeling. It is based on Sugeno-Fuzzy inference system whereby the output of the system is only either constant or linear. This

ISBN: 978-1-941968-02-4

technique uses if-then rules in the learning process.

ANFIS model consists of five layers as shown in the Figure 1. It is somewhat similar to Fuzzy Inference System in terms of the system components; however it consists of additional neural network component blocks which make it useful for prediction systems. The five layers of ANFIS model are namely: input layer, membership function layer, rule layer, norm layer, and output layer respectively [6]. The five layers of ANFIS system are described below.

Layer 1: This layer is known as input layer. It consists of input variables. This layer sends the input values to the next layer.

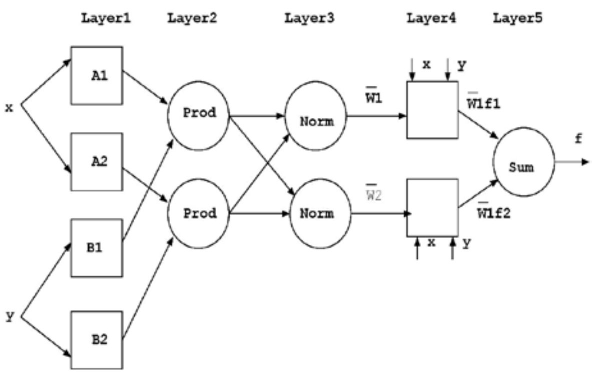

Figure 2. Representation of the five layers of ANFIS [6]

Layer 2: All the nodes in this layer are fixed and as named as Prod. It presents the output as a product of all the signals.

Layer 3: The pre-condition matching of the fuzzy rules are carried out in this layer.

Layer 4: This layer is responsible to assign output values from the inference of rules. This layer is also known as defuzzification layer.

Layer 5: This layer is the output layer which basically sums up all the products from the defuzzification layer.

3.2 System Details

The system model is constructed to classify the occurrence of flare at a point of time in the future. Firstly, the solar flare occurrence data was collected from NGDC website which is made publicly available. As it can be seen in Figure 1, the temperature and date and exact time of occurrence are available in the raw dataset. In order to make the data as sequential as possible, only the date of flare occurrence and the type of flare was chosen to be used in the design. The dataset was arranged on daily basis so that the system can predict the occurrence of flare. On the other hand, the temperature has not been chosen because it not provided for all flares.

In the dataset each date was taken as one sample with the corresponding flare. It is important to note that for training and testing purposes, one thousand data-pairs are used. Out of which, half of the pairs are assigned for training the system, while the other half of the pairs are used for validating the performance of the proposed system.

ANFIS can only be applied in Sugeno-type system in which for each input has a single output which is a constant. The system favors a total of 16 membership functions. At the beginning of the simulation the original data was plotted using linear graph. This is to provide a reference point and compare the predicted results.

3.3 Process Flow of the System

The data expresses the time and type of flare occurred in the past. A flow chart explaining the working process of the solar flare prediction system is shown in the Figure 3.

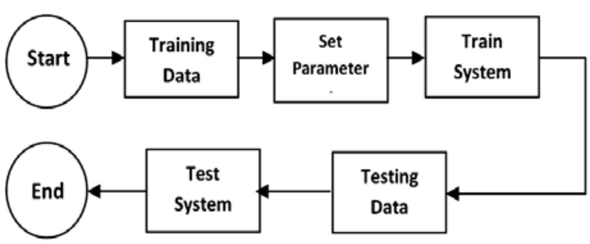

Figure 3. Process flow of the system.

At the first stage, the data are set according to the requirement of the ANFIS. This process includes assignment of the number of samples

ISBN: 978-1-941968-02-4 ©2014 SDIWC

as well as the number of inputs use in the system. Since ANFIS in Sugeno-type, it can only have a constant or linear output. Thus, it is important to determine the nature of the output at this stage. In the next stage, the membership functions for each input, and the number of data use in each input are determined.

The system is trained using the training dataset. It is important to know that the training and testing datasets used in the system are totally different sets of data pairs to obtain valid results. The performance of the system depends on the training process, thus to achieve better results. Training process is optimized by varying parameters like the number of epochs, and the type of membership functions. Once training is successful, the system is trained and thus can be used for validating the data.

When checking the performance of the system, it uses the inputs from the checking data and predicts the output. The output of the checking dataset is then used to verify the predicted data for the given inputs. The prediction of the system using the testing dataset is plotted so that it can be compared with the original time series data.

4 RESULTS AND DISCUSSIONS

The prediction results are compared with the original data to conclude the accuracy of system performance. This has been done by calculating the error between the training data and the testing data is determined. The results are plotted in terms of the flare prediction at least 24 hours before its occurrence time and the accuracy of correct flare occurrence. The accuracy of flare prediction occurrence determines if the system is able to predict the occurrence of a solar flare irrespective of flare class. On the other hand, the flare classification prediction determines if the system is capable of predicting the exact flare which is going to occur in future.

The proposed system for solar flare prediction needs to be validated in order to determine its performance. Hence three test criteria have been chosen that include Root Mean Squared Error (RMSE), Prediction Validation, and Percentage Error.

4.1 Root Mean Square Error (RMSE)

The RMSE is a commonly used validation technique in prediction systems. In this technique, the error is calculated by taking the square root of average of the square of the total error. It calculates the difference between the actual and the predicted values from the proposed system.

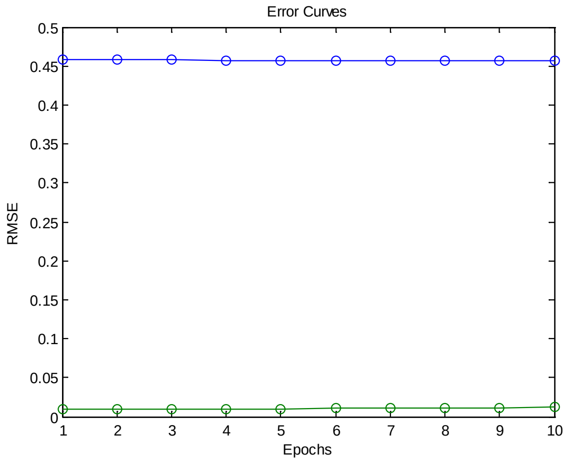

Figure 4. Training and Checking RMSE errors

Figure 4 shows the training and the checking RMSE errors.

Table 1: Training and Checking error

Trn_error	Chk_error
0.4579	0.0100
0.4577	0.0100
0.4576	0.0100
0.4575	0.0100
0.4573	0.0100
0.4572	0.0110
0.4570	0.0110
0.4569	0.0110
0.4567	0.0110
0.4565	0.0121

ISBN: 978-1-941968-02-4

The errors valuse are tabulated in Table 1 as well for the simplicity. It can be observed that the training error significantly higher than the checking error. The average of RMSE error is about 1%.

4.2 Accuracy of Occurrence of Flare Prediction

The main objective of the solar flare prediction systems is the correct prediction of flare occurrence and accuracy test is one of the important tests in validating such systems. The prediction accuracy is evaluated by comparing the predicted results with the original data. The proposed system has been tested for flare occurrence prediction, and accuracy of correct type of flare prediction. The accuracy of flare occurrence prediction determines if a flare is about to occur in the near future means at least within 24 hours after the prediction, despite of not being able to determine the exact flaring type. The accuracy of correct flare type prediction has also been carried out to test if the system is able to correctly predict the type of flare about to occur.

Table 2: Accuracy of flare prediction

	No Flare	Flare
Actual	71	929
Predicted	34	794

The prediction accuracy of system is determined by calculating the difference of the original data, and system predicted data. Table 2 provides the original data for testing purpose. It is noticed that there are 929 data pairs have a flare and 71 data pairs are with no flares. A total of 71 data pairs were used for training purpose.

The results show that 34 data pairs, 47.88 % of the testing data, have correctly been predicted that there will be no flare.

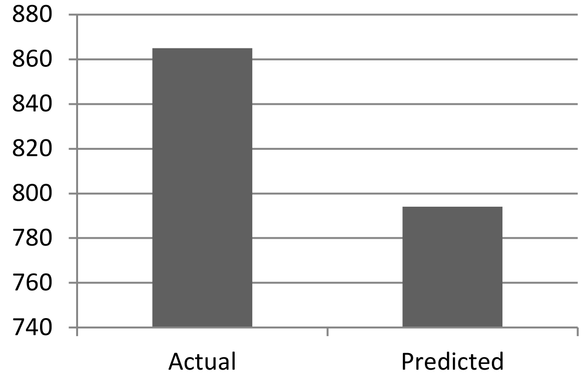

Figure 5. Total flares vs predicted flares.

On the other hand, Figure 5 shows that 929 cases of solar flare occurrence were presented in the dataset used in the simulation in which the system was able to predict 794 cases accurately. This is exactly 85.47% correct flare occurrence prediction

Figure 6 illustrates the percentage of flare occurrence predicted by the proposed model. The “No Flare” bar represents the percentage of how accurately the system predicted if there is no solar flare which is 3.4 %, while the “Flare” bar represents the percentage of the prediction of solar flare occurrence and it is 79.4 %.

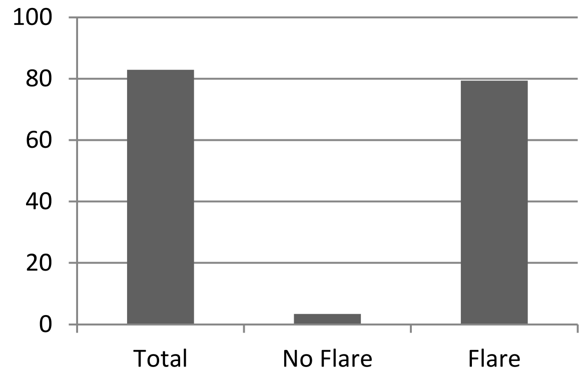

Figure 6. Prediction Accuracy

The “Total” bar represents the accuracy of total correct solar flare predictions over the total data pairs used in the proposed system which is 82.8 %. The results show that the system has 82.8 % accuracy.

4.3 Accuracy of Flare Classification Prediction

Table 4 provides the breakdown of the solar flare data pairs. There are a total of 929 data pairs. Out of which 865 are the M-class flares data and 64 data pairs are X-class flares.

Table 4: Flare type prediction

	No Flare	M-Flare	X-Flare
Actual	71	865	64
Predicted	34	786	8

Figure 7 demonstrates the M-Flare predicted bar graph. The “Actual” bar shows that there are 865 M class flare data and “Predicted” bar shows that 786 cases are correctly predicted by the proposed system.

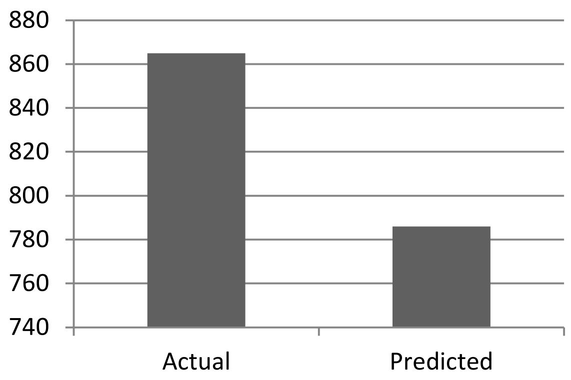

Figure 7. M-Class flare prediction.

Figure 8 illustrates the X-Class flare prediction bar graph. A total of 64 X-Class data flares were used. Out of which only 8 cases were correctly predicted by the system.

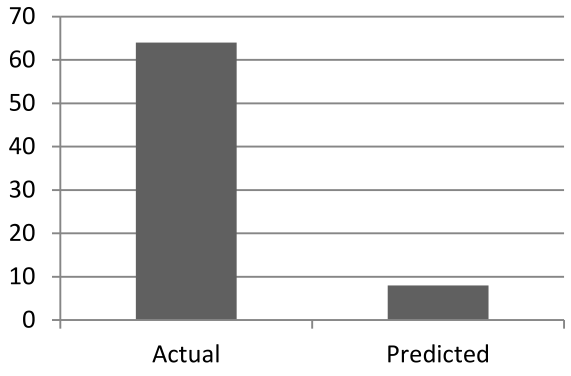

Figure 8. X-Class flare prediction

Figure 9 shows the overall flare prediction accuracy graph. From the graph, it can be observed that the number of correct flare obtained from the system is comparatively close to that of the number of actual flare provided to the system.

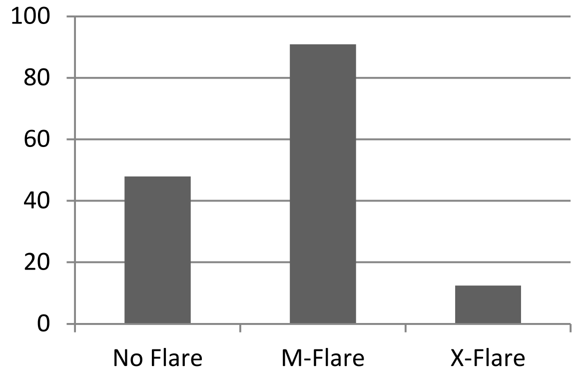

Figure 9. Overall Flare Prediction accuracy

The system is capable of predicting 90.86% of M type flares; however the percentage of correct prediction of X-Class flares is significantly very small which is 12.5%.

5 CONCLUSION

Solar flare is one of the major threats to satellites revolving in their orbits around the Earth. An intelligent ANFIS based flare prediction system is proposed. The performance of the proposed system is analyzed based on the accuracy of prediction of possible solar flare occurrence, prediction of solar flare class, and the error in the prediction. The results show that the system has 82.8% of accuracy for prediction of solar flare occurrence, while 79.4% in case of flare class prediction. The system has M-Class flare prediction accuracy of 90.86 %.

REFERENCES

[1] D. Stillman, “What Is a Satellite?,” Available at http://www.nasa.gov, 2010.

[2] S. P. Wyatt, “The Electrostatic Charge of Interplanetary Grains,” Planetary and Space Science, 1963.

[3] S. Guyer, and Z. Can. “Solar Flare Effects on the Ionosphere,” 6th International Conference in

Recent Advances in Space Technologies (RAST), 2013, Istanbul, pp. 729–733.

[4] R Qahwaji. and T. Colak, “Neural Network-based Prediction of Solar Activities,” Conference on Cybernetics and Information Technologies, Systems and Applications (CITSA2006) 2006, Orlando.

[5] X. Zhang, “Image Feature Extraction for Solar Flare Prediction,” 4th International Congress on Image and Signal Processing, 2011. Shanghai, pp. 910–914.

[6] A. Kusagur, S. F. Kodad, and R. B. V. Sankar, “Modeling, Design & Simulation of an Adaptive Neuro-Fuzzy Inference System (ANFIS) for Speed Control of Induction Motor,” International Journal of Computer Applications , Volume 6– No.12, 2010, pp. 29-44, September 2010.